PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Usefulness of Applying Research Reporting Guidelines as Writing Aid Software: a crossover randomized controlled trial
AUTHORS	Hawwash, Dana; Sharp, Melissa K; Argaw, Alemayehu; Kolsteren, Patrick; Lachat, Carl

VERSION 1 - REVIEW

REVIEWER	Elizabeth Considine University College Cork Ireland
REVIEW RETURNED	16-May-2019

GENERAL COMMENTS	This paper requires specialist statistical review which I am unqualified to do. I recommend that this be carried out by another reviewer. Page 2: Mention the primary outcome "intention of use" in the objectives section of the abstract, only secondary objectives are mentioned Page 3: Results vs Page 17 "We found no significant difference in reported intention of use" vs line 478..."more than two thirds selected the writing aid as the preferred method of use for later use" There is a discrepancy in these two sentences. Please resolve. Point 2: Page3 Mention in the limitations box: the limitation of the study participants being mostly PHD students unfamiliar with reporting guidelines and their usage and elaborate how this would affect the study Page 6: Line 158. Mention the EQUATOR Network, where these 400 reporting guidelines can be found Page 6: Line 163 Rephrase as context is unclear... Example "Present use of reporting guidelines requires consideration, as research on the usefulness of reporting guidelines from the point of view of the authors is scarce" Page 7. Introduction /Objectives

Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

Page 7. Line 174
Insert "adherence" or "usage" after the word guideline

Page 7. Line 174 to Line 180.
The section from "For example…" to "…submission application" needs to be rewritten as it is grammatically incorrect and its meaning is therefore unclear and ambiguous.
Page 7 Line 181.
Change the word "none" to "neither". Also at this point in the manuscript it is worth mentioning that there does exist a recent authoring tool developed with the aim of increasing adherence to reporting guidelines at the time of writing.
https://www.mdpi.com/2218-1989/9/3/43A Tool to Encourage Minimum Reporting Guideline Uptake for Data Analysis in Metabolomics

Page 9. Line 256.
Explain why the flow charts were not included

Page 10. Line 290.
Change "In case" to "If"

Page 11. Line 325-Line326
Rephrase the section "for both….future use" as it is currently grammatically incorrect and I'm not clear on what the author is trying to say.

Page 12. Line 354.
Change "aim" to "aims" and change "co-author" to "co-authors"

Page 13. Line 380- Line382
"1) subjects' better PEU…." "2) subjects' better perceived ease of use" ….is this not the same point made twice?

Page 17. Line 484.
This line needs to be changed… In its current format it is too definite and absolute. I recommend changing to
"This study aimed to test
"This study attempted to test

Page 17. Line 486
Seems to be a word or phrase missing in the middle of the sentence.. to test what exactly
"This paper adds value to prior knowledge by using an intervention to test…" at this point the authors need to insert either "adherence to" /"uptake of" / "perceived ease of use of" or some other phrase before "reporting guidelines in a real writing process"

Page 18. Line 496.
Remove the word "use"

Page 18. Line 501.
Change "apply" to "application of"

Page 18. Line 521.
Consider changing "in a real life setting" to "an approximation of a real life setting"

Page 19. Line 545.
An important limitation of the study that the authors did not mention is the fact that 80% of participants were PhD students and therefore the selected participants in the study were not a balanced representative of typical authors, and half of participants had never used a reporting guideline before. Elaborate what effect this would have on the study.

Page 19. Line 547-549
Insert some extra text for better readability..
"This will offer the possibility to explore further the potential of the applicability of this Writing Aid to enhance reporting guideline adherence."

Page 19. Line 549-550
Consider inserting the following
"The Writing aid deserves attention for further testing in a broader set of participants more representative of typical authors and also product development."

Specific review checklist:
1. Is the research question or study objective clearly defined?
Page 7. Introduction /Objectives
Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

2. Is the abstract accurate, balanced and complete?

Page 2:
Mention the primary outcome "intention of use" in the objectives section of the abstract, only secondary objectives are mentioned

3. Are the outcomes clearly defined?
Page 7. Introduction /Objectives
Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

4. Are the study limitations discussed adequately?

Mention in the limitations box: the limitation of the study participants being mostly PHD students unfamiliar with reporting guidelines and their usage and elaborate how this would affect the study

7. If statistics are used are they appropriate and described fully?

This paper requires specialist statistical review which I am unqualified to do. I recommend that this be carried out by another reviewer.

REVIEWER	Vedran Katavic
	University of Zagreb School of Medicine, Croatia
REVIEW RETURNED	28-Jun-2019

GENERAL COMMENTS	The study of the intention to use, perceived usefulness, and ease of use of a Writing aid software which integrates four research reporting guidelines is well thought-through analysis of ways to improve/enhance adherence to appropriate reporting.
	The study is described in enough detail for it to be replicated. The results are presented in a clear and reasonable way, and the discussion section adequately draws on the results and interprets them well.
	The strength of the study is it being first to test the application of reporting guidelines in a real-life setting, in a diverse group of participants from various scientific and cultural backgrounds. The weaknesses of the study - relatively small sample size, short exposure and subjective nature of outcomes are disclosed and discussed well.
	The language used is of high quality. The appendices are very useful (and give full disclosure on the IC forms and information sheets). provide the full trial protocol and the manual to the software used.
	Interesting work!

REVIEWER	Wei Gao
	King's College London, UK
REVIEW RETURNED	21-Aug-2019

GENERAL COMMENTS	This paper presented the evaluation of a potentially useful tool for scientific writing – the Writing Aid software. Based on their study objectives, I am not sure if the study design is an appropriate one. Perhaps a qualitative interview would be better to answer the question. Please justify the choice of the RCT study design.
	Objectives: clear
	Abstract: concise but need some clarifications, e.g. design – unblinding; type of outcome measures - continuous or categorical, score range; statistical methods – why SEM.
	Introduction
	Overall well justified but can be made more concise. E.g. subheading "Objectives" can be removed and the paragraph can be shortened.
	Methods
	Participants selection process needs to be clearer. What is the potential selection bias (response rate)? There is no sample size planning description. The statistical method also needs justifications. For example, why t-test? What's the data distribution? Why SEM? How many variables were used in the SEM? Was the SEM considered in the study planning stage? i.e. sample size needed for the SEM.
	Tables & Figures

	Relevant and clear. Perhaps to drop Table 3 and integrate the contents to Figure 1.

REVIEWER	Gordon S. Doig University of Sydney, Australia
REVIEW RETURNED	21-Aug-2019

| GENERAL COMMENTS | Review of manuscript bmjopen-2019-030943 titled "Usefulness of Applying Research Reporting Guidelines as Writing Aid Software: a crossover randomized controlled trial"

1. Please do not overuse novel acronyms; it makes your Abstract and Manuscript more difficult to read. I recommend replacing all novel acronyms (PU, PEU etc) with complete descriptions (Perceived Usefulness, Perceived Ease of Use. Also, if you intend to capitalise these constructs, I recommend capitalising your primary outcome: Intention of Use.
2. Abstract: Please remove the statement regarding 'mediation' of Intention to Use from the Abstract unless Perceived Ease of Use was ascertained BEFORE Intention to Use, in which case please report covariate adjusted p-value for Intention to Use as you would any other model adjusted for a baseline covariate.
3. Abstract, Conclusions: Focus your Conclusions on your Primary Outcome. It is important for your Reader to understand that your primary outcome was not significant.
4. Objectives, Page 7. Your primary outcome was clearly stated as 'Intention to Use'. Your sentence at Line 188 is misleading. Please alter it to read "and assessed participants intention to use it during the writing process."
5. Based on a read of your Methods, it appears Intention to Use, Perceived Ease of Use and Perceived Usefulness were all collected at the same time, using different instruments. Please remove the comments on Mediation from your Abstract, as it is clear this is pure hypothesis generation.
6. Page 15, Line 433-435. Please make it explicitly clear that your primary outcome was not statistically significant. Your current sentence 'implies' that 'a moderate effect was found' and that the Writing Aid 'seemed to perform better', This is unacceptable. Please alter your manuscript throughout to explicitly report that 'no effect was found' and that the Writing Aid was not significantly better than the Word Document with regards to your primary outcome.
7. Comments regarding Size and Strength of Effect: Line 371 reports categorisation of 'Size and stregth' using a measure of the difference / SD. Table 2 reports Effect Size for your primary outcome as 0.45. Absolute difference between 5.84 and 5.51 is 0.33. With a pooled SD of approximately 1.24, your difference / SD should be approximately 0.266. Likewise, the effect size for Perceived Ease of Use appears to be even smaller (.73/1.3=0.56). In Table 2, please report the absolute mean difference between each group and a P-value obtained from a t-test assessing mean difference (always report the metric assessed by the statistical test with the p-value). Next, report the SD around the mean difference.
8. Please adjust your Discussion to focus on your Primary outcome, and results presented as requested in Comment 7. |
|---|---|

Reviewer: 1

Reviewer Name: Elizabeth Considine

Institution and Country: University College Cork, Ireland

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This paper requires specialist statistical review which I am unqualified to do. I recommend that this be carried out by another reviewer.

Page 2:

Mention the primary outcome "intention of use" in the objectives section of the abstract, only secondary objectives are mentioned

ANS: The abstract has been revised accordingly Line 26-30 read "Objectives: To assess the intention of using a Writing Aid software, which integrates four research reporting guidelines (CONSORT, PRISMA, STROBE and STROBE-nut) and their Elaboration & Explanation (E&E) documents during the write-up of research in Microsoft Word compared to current practices."

Page 3: Results vs Page 17 "We found no significant difference in reported intention of use" vs line 478..."more than two thirds selected the writing aid as the preferred method of use for later use" There is a discrepancy in these two sentences. Please resolve.

ANS: A sentence in the text was added to clarify that preference might not lead to intention of use and actual use. The main outcome was intention of use, while the method of preference was an added question to give more context. Lines 527-530 read as follow "However, after being exposed to the two tools, more than two thirds of the sample answered that the Writing Aid was their preferred method of use. It is important to note that preferences might not lead to intention of use and actual use."

Point 2: Page3

Mention in the limitations box: the limitation of the study participants being mostly PHD students unfamiliar with reporting guidelines and their usage and elaborate how this would affect the study

ANS: Thank you for your valid comment. The limitations box can only have 5 bullets maximum according to BMJ open instructions to author. We argue that the bullets presented are all substantial and we would rather not substitute any of them. However, we have added this limitation in the discussion section, which now includes an elaboration on how this could have affected the study findings in line 576-579 "Fourth, purposive sampling was used. The majority (80%) of the participants were PhD students, which might be unrepresentative for other authors. Further assessment in authors with more seniority is required."

Page 6: Line 158. Mention the EQUATOR Network, where these 400 reporting guidelines can be found

ANS: line 154 which is revised to include the EQUATOR network

"Currently, there are more than 400 reporting guidelines indexed by the EQUATOR Network, an international organization that promotes the use of reporting guidelines"

Page 6: Line 163 Rephrase, as context is unclear... Example "Present use of reporting guidelines requires consideration, as research on the usefulness of reporting guidelines from the point of view of the authors is scarce"

ANS: Paragraph has been rephrased to bring clarity to context. The paragraph from line 157- 162 now reads as follows "However, present use of reporting guidelines requires consideration. Reporting guidelines are typically applied at the final stages of the writing process to address journal requirements 1. As a result, reporting guidelines might be considered as an administrative burden rather than a tool to improve research quality. Moreover, research on the usefulness of reporting guidelines from the point of view of the authors is scarce. "

Page 7. Introduction /Objectives Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

ANS: The intention of use was defined later in the methods section. The section has been moved to the introduction section to add clarity. The introduction section starting line 185-219 now reads:

"In recognition of these issues, we developed a Writing Aid tool that integrates the reporting guidelines and their Elaboration & Explanation documents, in the form of an Add-in for Microsoft Word (Version 1.0, Automaticals Consulting)2, and assessed participants intention to use it during the writing process using the Technology Acceptance Model (TAM). The TAM model has been validated and applied previously to test software in similar settings. It has been used in office environment operations (e.g. text editor, voicemail), software application development (e.g., software maintenance tool), and core business process software (e.g., production control tools)3. The overall objective was to investigate researchers' intention of using the reporting guidelines as a Writing Aid in Word versus the traditional approach of a Word document and the Elaboration & Explanation document. Secondary objectives included perceived usefulness, and perceived ease of use. We also assessed how perceived usefulness and perceived ease of use were associated with the intention of use. The questionnaires contained questions with a 7-point Likert-scale responses ranging from extremely unlikely ('1') to extremely likely ('7') (Appendix 2). The intention of use outcome was constructed from two questions:

Q1- Assuming I have access to the reporting guidelines documents (as a MS Word table and elaboration and explanation document/ Writing Aid), I intend to use it

and

Q2- Given access to the reporting guidelines documents (as a MS word table and elaboration and explanation document/ Writing Aid), I predict that I would use it.

The perceived usefulness outcome was constructed from four questions. Participants were asked to rate the reporting guideline usefulness (using the Writing Aid and Word Document) based on the tool's ability to improve completeness of reporting, increase productivity, enhance effectiveness, and usefulness. The perceived ease of use outcome was also constructed from four questions. Participants were asked to rate the ease of using the reporting guidelines (using the Writing Aid and Word Document) based on how flexible, easy to use, easy to provide guidance, clear and understandable it was to interact with."

Page 7. Line 174 Insert "adherence" or "usage" after the word guideline

ANS: Done, line 174 includes adherence "There have been initiatives attempting to develop new tools and test their impact on reporting guideline adherence."

Page 7. Line 174 to Line 180.The section from "For example…" to "…submission application" needs to be rewritten as it is grammatically incorrect and its meaning is therefore unclear and ambiguous.

ANS: line 174-180 which is now line 174-183 has been revised and it reads now as follows

"There have been initiatives attempting to develop new tools and test their impact on reporting guideline adherence. For example, the online COBWEB tool4 5 guides authors on how to apply the CONSORT reporting guideline to their manuscript and Penelope, an automated online tool, generates automatic checks of manuscripts written in MS Word. Penelope is currently being integrated and tested in an online journal submission application6-8. A recent study also developed a writing tool, and a template with the minimum amount of information to report regarding data handling of biomarkers in metabolomics9. Nevertheless, none of these efforts focuses on the uptake of several reporting guidelines during the writing process, using common offline writing platforms such as Microsoft Word.

Page 7 Line 181. Change the word "none" to "neither". Also at this point in the manuscript it is worth mentioning that there does exist a recent authoring tool developed with the aim of increasing adherence to reporting guidelines at the time of writing.

https://www.mdpi.com/2218-1989/9/3/43A Tool to Encourage Minimum Reporting Guideline Uptake for Data Analysis in Metabolomics

ANS: The third tool suggested above was added to the previous initiatives; see answer to the previous comment. The text on line 181-183 now reads as follows: "Nevertheless, none of these efforts focuses on the uptake of several reporting guidelines during the writing process, using common offline writing platforms such as Microsoft Word."

Page 9. Line 256. Explain why the flow charts were not included

ANS: There are a few reasons why the flow charts were not included

1- As mentioned in the objectives, "The study was conceived by authors of the STROBE-nut as an approach to improve its uptake. However, during the study set up, and the software development, it became clear that the intention of use and the software developed are relevant for other reporting guidelines. As a proof of concept study, CONSORT, PRISMA, and STROBE were included to test the wider application within other research designs and fields." Neither STROBE nor STROBE nut have a flowchart included.

2- We were mainly focused on the writing process, and although a flow chart provides crucial information, its inclusion would not affect the writing process

3- Lastly, due to limited funding and technicality problems, we had to make certain choices to limit cost. The addition of graphs would have needed further sophisticated software programming.

The following lines (288-293) were included to add clarity to the text: "The decision was made as the study mainly focused on the writing process. Although flow charts provide crucial information for the manuscript, they are not typically part of the narrative sections. Moreover, their inclusion needed further sophisticated software programming, time and resources."

Page 10. Line 290. Change "In case" to "If"

ANS: "In case" has been changed to "if" (line 328)

Page 11. Line 325-Line326 Rephrase the section "for both….future use" as it is currently grammatically incorrect and I'm not clear on what the author is trying to say.

ANS: The section has been rephrased to include the actual questions as being asked during the study and has been moved into the introduction section while addressing your previous comment. "The intention of use outcome was constructed from two questions:

Q1- Assuming I have access to the reporting guidelines documents (as a MS Word table and elaboration and explanation document/ Writing Aid), I intend to use it

and

Q2- Given access to the reporting guidelines documents (as a MS word table and elaboration and explanation document/ Writing Aid), I predict that I would use it.

Page 12. Line 354. Change "aim" to "aims" and change "co-author" to "co-authors"

ANS: done "aim" was changed to "aims" and "co-author" was changed to "co-authors". lines 375-376)

Page 13. Line 380- Line382 "1) subjects' better PEU…." "2) subjects' better perceived ease of use" ….is this not the same point made twice?

ANS: The two points explains two different paths in the model. Point one refers to the immediate relation between perceived ease of use and intention of use while point two refers to a pathway between perceived ease of use and intention of use mediated by perceived usefulness (perceived ease of use → perceived usefulness→and intention of use).

The lines have been rephrased in the text on (line 403-408) to add clarity

1) Immediate pathway between subjects' better PEU for Writing Aid compared to the Word Document leading to a better intention of use, and 2) Chain pathway in which subjects' better perceived ease of use could lead to a better perceived usefulness for the Writing Aid than the Word Document and finally resulting in a better intention of use.

Page 17. Line 484. This line needs to be changed… In its current format it is too definite and absolute. I recommend changing to "This study aimed to test "This study attempted to test

ANS: Done wording has been changed to "this study attempted to test" on line 510

Page 17. Line 486 Seems to be a word or phrase missing in the middle of the sentence.. to test what exactly "This paper adds value to prior knowledge by using an intervention to test…" at this point the authors need to insert either "adherence to" /"uptake of" / "perceived ease of use of" or some other phrase before "reporting guidelines in a real writing process"

ANS: The phrase has been adjusted including the word uptake of to add clarity. The text now reads "This paper adds value to prior knowledge by using an intervention to test the uptake of reporting guidelines in a real writing process, using all sections of the paper." Line 513

Page 18. Line 496. Remove the word "use"

ANS: the word "use" was removed

Page 18. Line 501. Change "apply" to "application of"

ANS: Done, "apply" was changed to "application of " line 531

Page 18. Line 521. Consider changing "in a real life setting" to "an approximation of a real life setting"

ANS: Done, "in a real life setting" was changed to "an approximation of a real life setting" line 552-553

Page 19. Line 545. An important limitation of the study that the authors did not mention is the fact that 80% of participants were PhD students and therefore the selected participants in the study were not a balanced representative of typical authors, and half of participants had never used a reporting guideline before. Elaborate what effect this would have on the study.

ANS: A fourth limitation has been added to the text. The text now reads on line 576-579 "Fourth, purposive sampling was used. The majority (80%) of the participants were PhD students, which might be unrepresentative for other authors. Further assessment in authors with more seniority is required."

Page 19. Line 547-549 Insert some extra text for better readability. "This will offer the possibility to explore further the potential of the applicability of this Writing Aid to enhance reporting guideline adherence."

ANS: Done, extra text was added as suggested for better readability. Line 588-590 "This will offer the possibility to further explore the potential applicability of our Writing Aid to enhance reporting guideline adherence."

Page 19. Line 549-550 Consider inserting the following "The Writing aid deserves attention for further testing in a broader set of participants more representative of typical authors and also product development."

ANS: The text has been revised to include the recommended addition line 590-5972 "The findings of this study are encouraging for further product development and testing in a more representative sample of researchers."

Specific review checklist: 1. Is the research question or study objective clearly defined?

Page 7. Introduction /Objectives

Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

Addressed higher

2. Is the abstract accurate, balanced and complete?

Page 2:

Mention the primary outcome "intention of use" in the objectives section of the abstract, only secondary objectives are mentioned

Addressed higher

3. Are the outcomes clearly defined?

Page 7. Introduction /Objectives

Define "intention of use" clearly either in the introduction or the objectives section. For an example see Table 3 in the following publication. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511365/

Addressed higher

4. Are the study limitations discussed adequately?

Mention in the limitations box: the limitation of the study participants being mostly PHD students unfamiliar with reporting guidelines and their usage and elaborate how this would affect the study

Addressed higher

7. If statistics are used are they appropriate and described fully?

This paper requires specialist statistical review which I am unqualified to do. I recommend that this be carried out by another reviewer.

Reviewer: 2

Reviewer Name: Vedran Katavic

Institution and Country: University of Zagreb School of Medicine, Croatia

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

The study of the intention to use, perceived usefulness, and ease of use of a Writing aid software which integrates four research reporting guidelines is well thought-through analysis of ways to improve/enhance adherence to appropriate reporting. The study is described in enough detail for it to be replicated. The results are presented in a clear and reasonable way, and the discussion section adequately draws on the results and interprets them well. The strength of the study is it being first to test the application of reporting guidelines in a real-life setting, in a diverse group of participants from various scientific and cultural backgrounds. The weaknesses of the study - relatively small sample size, short exposure and subjective nature of outcomes are disclosed and discussed well. The language used is of high quality. The appendices are very useful (and give full disclosure on the IC forms and information sheets). provide the full trial protocol and the manual to the software used. Interesting work!

ANS: Dear reviewer we appreciate that you have read our work with great interest.

Reviewer: 3

Reviewer Name: Wei Gao

Institution and Country: King's College London, UK

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This paper presented the evaluation of a potentially useful tool for scientific writing – the Writing Aid software. Based on their study objectives, I am not sure if the study design is an appropriate one. Perhaps a qualitative interview would be better to answer the question. Please justify the choice of the RCT study design.

ANS: Thank you for this valuable comment. Indeed the study objective is to study the potential usefulness of the writing aid tool for scientific writing. A qualitative interview could give us insight on the usefulness of the writing aid tool but will not provide estimates on how it compares to the traditional way of addressing reporting guidelines. We chose a randomized controlled trial, since it is the most reliable method to test new intervention against established/ existing methods and measure the difference. We further chose a crossover design to minimize confounding covariates as each participant serves as his/her own control.

Objectives: clear

ANS: Thank you

Abstract: concise but need some clarifications, e.g. design – unblinding; type of outcome measures - continuous or categorical, score range; statistical methods – why SEM.

ANS:

• Design is clarified and non blinding is included

Design: Two- arms crossover randomised controlled trial with no blinding and no washout period.

• Type of the outcome measures, score range, statistical methods and why SEM are now included in the abstract

Primary and secondary outcomes: Using the Technology Acceptance Model, we assessed the primary outcome: the difference in the mean of intention of use; and secondary outcomes: the difference in mean perceived ease of use and perceived usefulness. The three outcomes were

measured using questions with a 7-point Likert-scale. Structural Equation Modeling (SEM) was applied to explore the relationships between the outcomes.

Introduction

Overall well justified but can be made more concise. E.g. subheading "Objectives" can be removed and the paragraph can be shortened.

ANS: the sub-heading "Objective " was removed. We have attempted to make the text more concise, yet due to another reviewer's comment on the need to clarify the outcome measures, new elements have been added to the text. A native English researcher read the text and helped us to rewrite it more concisely.

Methods

Participants selection process needs to be clearer. What is the potential selection bias (response rate)? There is no sample size planning description. The statistical method also needs justifications. For example, why t-test? What's the data distribution? Why SEM? How many variables were used in the SEM? Was the SEM considered in the study planning stage? i.e. sample size needed for the SEM.

ANS: Below are detailed answers to your inquires. We note this information was previously included in the protocol and have now moved this to the manuscript to clarify where necessary.

• Selection Process

Participant selection process is explained in the protocol https://biblio.ugent.be/publication/8559624. We have revised the text in the paper to include more clarification: the text reads as follows now: The discussion has more clarification on bias that can be introduced by purposive sampling. Line 255-264.

"Purposive sampling and snowballing was used to recruit participants from May until the end of October 2018. Eligible subjects were doctoral and post-doctoral students who were writing or recently had published a paper in any biomedical research field in the previous six months. Personalized email invitations were disseminated to potentially eligible students at Ghent University, the University of Split, the Methods in Research on Research "MiRoR" network, and at conferences (Federation of European Nutrition Societies, Belgrade 2018, Tropentag Gent 2018, The Cochrane Colloquium 2018). Twitter and posters were also used to circulate the invitation to a wider audience10."

• What is the potential selection bias (response rate)?

The potential selection bias is now added at the end in the discussion section as a limitation to the study. It was difficult to track response rate, as recruitment methods used network snowballing, social media, and was spread in conferences, on twitter, and as posters at different faculties. Many PhD students, who received the invitation, work on other topics using other designs. However, in this study only those who willingly wanted to participate n=54 completed the study. The bias is discussed in the limitation section. This is now added in the text: line 438-440 "It was not possible to assess response rate, as recruitment methods used a snowballing approach. However, in this study only those who willingly wanted to participate n=54 completed the study."

Further addition in the discussion on the potential selection bias is now included in lines 576-579 "Fourth, purposive sampling was used. The majority (80%) of the participants were PhD students, which might be unrepresentative for other authors. Further assessment in authors with more seniority is required."

• Sample size calculation

Estimating a sample size was not tenable as there are no previous studies that reported an effect size. This has been reported in the protocol and on line 250 of the text

"No formal sample size calculation was conducted and we aimed to collect as many responses as possible."

The statistical method also needs justifications.

• For example, why t-test? What's the data distribution?

T-test was used, after the confirmation of the normality of the data by using histograms. The mean difference of intention of use, perceived ease of use and perceived usefulness were all normally distributed. This is now added in the results section line 390-392.

"We have used t-test to test the difference in mean intention of use, perceived usefulness and ease of use after confirming normality of data."

• Why SEM?

Structural Equation Modeling was used as a secondary analysis to test the pathways used in the Technology Acceptance Model11. This was explained in the study protocol "Structural equation modelling (SEM) will be used to assess the associations in the technology acceptance model. In addition, mediating and moderating analyses will be conducted to provide more insight into intervention effects." We chose to analyse the data using structural equation modeling based on litrature11.

• How many variables were used in the SEM? Was the SEM considered in the study planning stage? i.e. sample size needed for the SEM.

28 variables were used in the model. Yes, SEM was considered during the study-planning phase, and was announced in the protocol https://biblio.ugent.be/publication/8559624. SEM was used as a secondary analysis to give further indication on the intention of use through perceived usefulness and ease of use. We acknowledge the discussion on the sample size needed for SEM12, and are aware that 54 might be underpowered for this particular analysis. This is now added to the limitation section of the discussion on lines 567-570

"SEM, which was conducted as secondary analysis, was potentially underpowered. A larger sample size could have increased the power of the study, the statistical significance, and the bias in the parameter estimates used in the SEM12"

• Tables & Figures

Relevant and clear. Perhaps to drop Table 3 and integrate the contents to Figure 1.

ANS: We appreciate the comment that the tables and figures are relevant and clear. After discussions between the co-authors, we decided to keep Figure 1 and Table 3 separate since the two items explain and communicate distinct ideas. Figure 1 explains the methodology and Table 3 elaborates on the results, direct, indirect and the total effect. This would leave the figure very crowded with information

Reviewer: 4

Reviewer Name: Gordon S. Doig

Institution and Country: University of Sydney, Australia

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

Review of manuscript bmjopen-2019-030943 titled "Usefulness of Applying Research Reporting Guidelines as Writing Aid Software: a crossover randomized controlled trial"

1. Please do not overuse novel acronyms; it makes your Abstract and Manuscript more difficult to read. I recommend replacing all novel acronyms (PU, PEU etc) with complete descriptions (Perceived Usefulness, Perceived Ease of Use. Also, if you intend to capitalise these constructs, I recommend capitalising your primary outcome: Intention of Use.

ANS: The whole text has been updated without the use of novel acronyms including PU and PEU.

2. Abstract: Please remove the statement regarding 'mediation' of Intention to Use from the Abstract unless Perceived Ease of Use was ascertained BEFORE Intention to Use, in which case please report covariate adjusted p-value for Intention to Use as you would any other model adjusted for a baseline covariate.

ANS: We acknowledge that the use of the term "mediation" creates some confusion. We adjusted the language to describe what was done; structural equation modeling to explore if the technology acceptance model can explain further the finings in our study. We followed the analysis as pre-

specified in the protocol https://biblio.ugent.be/publication/8559624. We chose to analyse the data using structural equation modeling, based on litrature11. The abstract now reads as follows "Secondary analysis using Structural Equation Modeling (SEM) was applied to explore the relationships between the outcomes." &" In the SEM analysis, participants' intention of using the tools was indirectly affected by perceived ease of use (beta 0.5 P = 0.02)."

We have also rephrased the text on line 397-400 "To provide an explanation for participants' intention of using the tools (Writing Aid vs. Word Document) that is related to perceived ease of use and/or perceived usefulness, we also conducted Structural Equation Modeling (SEM) guided by the Technology Acceptance Model (TAM) (Figure 1)."

We have also included a limitation regarding the lack of temporal order in assessing the intention of use, perceived ease and usefulness, in the discussion line 579-582 "Fifth, intention of use, perceived ease of use and perceived usefulness were all collected at the same time, thus not allowing enough time for participants to experiment with the tools and assess the intention of use and actual use correctly".

3. Abstract, Conclusions: Focus your Conclusions on your Primary Outcome. It is important for your Reader to understand that your primary outcome was not significant.

ANS: We have updated the abstract conclusion to include the following: Results: No significant difference in reported intention of use (Mean difference and 95% CI: 0.25 (-0.05 - 0.55), P = 0.10), &"Conclusions: Despite no significant difference in the intention of use between the tools, administering reporting guidelines as Writing Aid is perceived as easier to use, offering a possibility to further explore its applicability to enhance reporting adherence."

4. Objectives, Page 7. Your primary outcome was clearly stated as 'Intention to Use'. Your sentence at Line 188 is misleading. Please alter it to read "and assessed participants intention to use it during the writing process."

ANS: The primary outcome of this study is indeed the intention of use. Line 188 has been altered from to "…and assessed participants intention to use it during the writing process."

5. Based on a read of your Methods, it appears Intention to Use, Perceived Ease of Use and Perceived Usefulness were all collected at the same time, using different instruments. Please remove the comments on Mediation from your Abstract, as it is clear this is pure hypothesis generation.

ANS: The intention to use, perceived ease of use and perceived usefulness were all assessed at the same time, using different instruments. Given the exploratory nature of the SEM, the conclusion

requires indeed further confirmation. We agree to lessen the attention on this conclusion, and keep it clearly described within the paper as secondary analysis. The abstract now reads as follows "Secondary analysis using Structural Equation Modeling (SEM) was applied to explore the relationships between the outcomes."

We have also included a limitation regarding this in the discussion line 579-582 "Fifth, intention of use, perceived ease of use and perceived usefulness were all collected at the same time, thus not allowing enough time for participants to experiment with the tools and assess the intention of use and actual use correctly".

6. Page 15, Line 433-435. Please make it explicitly clear that your primary outcome was not statistically significant. Your current sentence 'implies' that 'a moderate effect was found' and that the Writing Aid 'seemed to perform better', This is unacceptable. Please alter your manuscript throughout to explicitly report that 'no effect was found' and that the Writing Aid was not significantly better than the Word Document with regards to your primary outcome.

ANS: We agree, the way the text written can be misleading thus the text has been altered and the findings are explicitly clear. The text on lines 459-463 now reads" Table 2 shows that there is no significant difference in the performance between the Writing aid and the Word Document for both the primary outcome; intention of use (Mean difference and 95% CI: 0.25 (-0.05 - 0.55), P = 0.10), and perceived usefulness (Mean difference and 95% CI: 0.19 (-0.035-0.41), P = 0.10). A significant effect was found when comparing the perceived ease of use of the Writing Aid compared to the Word Document (Mean difference and 95% CI: 0.59 (0.29-0.89), P < 0.001)."

7. Comments regarding Size and Strength of Effect: Line 371 reports categorisation of 'Size and stregth' using a measure of the difference / SD. Table 2 reports Effect Size for your primary outcome as 0.45. Absolute difference between 5.84 and 5.51 is 0.33. With a pooled SD of approximately 1.24, your difference / SD should be approximately 0.266. Likewise, the effect size for Perceived Ease of Use appears to be even smaller (.73/1.3=0.56). In Table 2, please report the absolute mean difference between each group and a P-value obtained from a t-test assessing mean difference (always report the metric assessed by the statistical test with the p-value). Next, report the SD around the mean difference.

ANS: We redid the analysis, and corrected the error. We have done the analysis adjusting for period effect, thus the final result should be divided by 2 as recommended by Hills M, & Armitagen P.13 We have included the corrected absolute mean difference and SD in the modified table including the corresponding effect size, which now read as 0.23 for the primary outcome intention of use, and 0.54 for perceived ease of use (table below). The text has been modified accordingly. We appreciate the comment and we are pleased we could revise the findings before submission.

Table 2 Effect of the intervention on primary and secondary outcomes, Mean (SD), comparing Writing Aid and MS word tools in a crossover design (n = 54)

Outcomes (factor score) MS word

Mean (SD) Writing Aid

Mean (SD) Mean difference and 95% CI1 P value of mean difference Effect-size 95% CI1

Intention of use 5.51 (1.24) 5.84 (1.24) 0.25 (-0.05 - 0.55) P = 0.10 0.23 (-0.05- 0.5)

Perceived usefulness 5.38 (1.14) 5.63 (1.06)

0.19 (-0.035-0.41) P = 0.10 0.23 (-0.04-0.5)

Perceived ease of use 5.25 (1.30)

5.98 (0.93) 0.59 (0.29-0.89) P < 0.001 0.54 (0.25-0.83)

8. Please adjust your Discussion to focus on your Primary outcome, and results presented as requested in Comment 7.

ANS: The discussion has been adjusted to focus on the primary outcome and results presented as requested in Comment 7

Line 516 -522 "In the present study, participants indicated no significant difference in intention of use, and perceived usefulness between the two tools. This can be explained by the fact that the two applied interventions contained the same recommendations for reporting. However, participants perceived the Writing Aid to be easier to use than the Word Document with a significant effect. This can be attributed to the difference in application characteristics (integrated software vs. a MS Word document)."

References

1. Hopewell S, Altman DG, Moher D, et al. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal 'Instructions to Authors'. Trials 2008;9:20. doi: 10.1186/1745-6215-9-20

2. WritingAidTool [program]. 1 version. Github https://github.com/carllachat/WritingAidTool, 2016.

3. Legrisa P, Inghamb J, Collerette P. Why do people use information technology?

A critical review of the technology acceptance model. Information & Managemen 2001;40(2003):191-204.

4. Barnes C, Boutron I, Giraudeau B, et al. Impact of an online writing aid tool for writing a randomized trial report: the COBWEB (Consort-based WEB tool) randomized controlled trial. BMC Med 2015;13:221. doi: 10.1186/s12916-015-0460-y

5. COBWEB Consort-based WEB tool an online writing aid tool for writing a randomized trial report [Available from: https://cobweb.clinicalepidemio.fr/ accessed 3rd September 2019.

6. Penelope [cited 2019 21st January 2019]. Available from: https://www.penelope.ai accessed 2019.

7. Reporting checklists for medical researchers [cited 2019 21st January 2019]. Available from: https://www.goodreports.org2019.

8. Penelope. Case Study - BMJ Open 2018 [cited 2018 9th November]. Available from: https://www.penelope.ai/blog/2018/2/12/9id69afc4jd8sy6h36vc2x8px69myo accessed 9th November 2018.

9. Considine EC, Salek RM. A Tool to Encourage Minimum Reporting Guideline Uptake for Data Analysis in Metabolomics. Metabolites 2019;9(3) doi: 10.3390/metabo9030043

10. Twitter. 2018 [25th February 2019]. Available from: https://twitter.com/Danahawwash/status/1011238978471817216.

11. Gefen D, Straub DW, Boudreau M. Structural Equation Modeling and Regression: Guidelines for Research Practice. Communications of the Association for Information Systems 2000;4

12. Wolf EJ, Harrington KM, Clark SL, et al. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. Educ Psychol Meas 2013;76(6):913-34. doi: 10.1177/0013164413495237

13. Hills M, Armitagen P. The Two-Period Cross-Over Clinical Trial. Br J clin Pharmac 1979;8:7-20. doi: 0306-5251/79/070007-14 501.00

VERSION 2 – REVIEW

REVIEWER	Gordon S. Doig University of Sydney
REVIEW RETURNED	01-Oct-2019

GENERAL COMMENTS	The Editor needs to adjudicate whether this is a 'clinical trial', however, I am happy the authors have addressed my original comments. The WHO registry defines 'clinical trial' as "For the purposes of registration, a clinical trial is any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes." Personally and professionally, I do not believe this current paper falls under the WHO definition.